

IT IS CLAIMED:

1. An expression vector for expressing a multimeric polypeptide anchored on a surface of a genetically replicable package formed by a host, the expression vector comprising:

- 5 a vector segment encoding a polypeptide sequence having;
- i. a first polypeptide segment,
 - ii. a second polypeptide segment having therein a cleavable peptide sequence cleavable by a proteolytic agent, and,
 - 10 iii. a third polypeptide segment having therein an anchoring peptide sequence for anchoring said multimeric polypeptide to said surface of said genetically replicable package,
- the second polypeptide segment being between the first polypeptide segment and the third segment,
- whereby the cleavable peptide sequence is cleaved by the proteolytic agent and
- 15 whereby the first segment associates with the third segment to form the multimeric polypeptide.

2. The expression vector of claim 1, wherein the first and third polypeptide segments comprise an amino acid sequence derived from antibody light and heavy

20 chains.

3. The expression vector of claim 1, wherein the first and third polypeptide segments comprise the antigen binding regions of the variable domains of antibody light and heavy chains.

25

4. The expression vector of claim 1, wherein the first polypeptide segment comprises the variable domain and the constant domain of an antibody light chain, and the third polypeptide segment comprises the variable domain and a constant domain of the antibody heavy chain, such that when the first and third segments associate, the

30 product is a Fab antibody fragment.

5. The expression vector of claim 1, wherein the first polypeptide segment comprises the variable domain and the CH1 domain of an antibody heavy chain, and the

third polypeptide segment comprises the variable domain and the constant domain of the antibody light chain, such that when the first and third segments associate, the product is a Fab antibody fragment.

5 6. The expression vector of claim 1, wherein the first polypeptide segment comprises the variable domain and the constant domain of the antibody light chain, and the third polypeptide segment comprises the variable domain and the CH1 domain of an antibody heavy chain, such that when the first and third segments associate, the product is a Fab antibody fragment.

10

7. The expression vector of claim 1, wherein the first and third polypeptide segments comprise the variable domains of the light and heavy chains of a single antibody such that when the first and third segments associate, the product is an Fv antibody fragment.

15

8. The expression vector of claim 1, wherein the first polypeptide segment is N-terminal to the second polypeptide segment, and wherein the second polypeptide segment is N-terminal to the third polypeptide segment, and wherein the vector segment encoding the third polypeptide segment further includes one or more suppressable
20 nonsense codon(s) N-terminal to the anchoring segment.

9. The expression vector of claim 1, wherein the third polypeptide segment further includes a cleavable peptide sequence cleavable by a second proteolytic agent.

25 10. The expression vector of claim 9, wherein the first and second proteolytic agents are identical.

11. The expression vector of claim 1, wherein the proteolytic agent is selected from the group consisting of a chemical proteolytic agent and an enzymatic proteolytic
30 agent.

12. The expression vector of claim 1, wherein the proteolytic agent is expressed by the host.

13. The expression vector of claim 1, wherein the proteolytic agent is added such that it contacts and cleaves the second polypeptide segment.

5 14. The expression vector of claim 11, wherein the chemical proteolytic agent is an acid.

15. The expression vector of claim 1, wherein the cleavable peptide sequence comprises the sequence represented by SEQ ID NO:1.

10

16. The expression vector of claim 1, wherein the cleavable peptide sequence is not found in either the first or third polypeptide segments, and is recognized as a protein cleavage site by a proteolytic agent encountered in the host.

15 17. The expression vector of claim 1, wherein the polypeptide sequence further comprises one or more leader sequence(s) positioned upstream of the first polypeptide segment or third polypeptide segment or both first and third polypeptide segments.

18. The expression vector of claim 1, wherein the anchoring peptide comprises a segment encoding a phage coat protein.

20

19. The expression vector of claim 1, wherein the expression vector is selected from the group consisting of plasmids, phages, cosmids, phagemids, and viral vectors.

25 20. The expression vector of claim 1, wherein the expression vector is selected from the group consisting of M13, f1, fd, lf1, lke, Xf, Pf1, Pf3, λ , T4, T7, P2, P4, ϕ X-174, MS2 and f2.

21. The expression vector of claim 1, wherein the genetically replicable package is selected from the group consisting of a bacteriophage, a virus, a cell and a spore.

30

22. The expression vector of claim 21, wherein the cell is a bacterial cell.

23. The expression vector of claim 22, wherein the bacterial cell is selected from the group consisting of strains of *Escherichia coli*, *Salmonella typhimurium*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Neisseria gonorrhoeae*, and *Bacillus subtilis*.

5

24. The expression vector of claim 21, wherein the cell is a yeast cell.

25. The expression vector of claim 1, wherein the genetically replicable package is a filamentous bacteriophage specific for *Escherichia coli* and the anchoring peptide is a phage coat protein selected from the group consisting of coat protein III, coat protein pVI and coat protein VIII.

10

26. The expression vector of claim 25, wherein the filamentous bacteriophage is selected from the group consisting of M13 and fd.

15

27. The expression vector of claim 1, wherein the proteolytic agent is encoded by a nucleic acid sequence in the expression vector.

28. The expression vector of claim 1, wherein the proteolytic agent is encoded by a nucleic acid sequence in a second expression vector.

20

29. The expression vector of claim 1, wherein the cleavable peptide sequence comprises a disordered region cleavable by the proteolytic agent.

30. The expression vector of claim 1, wherein the cleavable peptide sequence comprises a specific peptide cleavage site cleavable by the proteolytic agent.

25

31. The expression vector of claim 1, wherein the cleavable peptide sequence includes a cleavage site for urokinase, pro-urokinase, thrombin, enterokinase, plasmin, plasminogen, TGF- β , staphylokinase, thrombin, Factor IXa, Factor Xa, a metalloproteinase, an interstitial collagenase, a gelatinase or a stromelysin.

30

32. The expression vector of claim 1, wherein the cleavable peptide sequence is cleavable by a protease selected from the group consisting of degP, degQ, degS and tsp.

5 33. The expression vector of claim 1, wherein the cleavable peptide sequence comprises a self-cleaving domain.

34. The expression vector of claim 33, wherein the self-cleaving domain is derived from an intein.

10

35. A host cell comprising the expression vector of claim 1.

36. The host cell of claim 35, wherein the proteolytic agent is a native proteolytic agent.

15

37. The host cell of claim 35, wherein the proteolytic agent is localized in the periplasm.

20

38. The host cell of claim 35, wherein the proteolytic agent is localized in the cytoplasm.

25

39. A method of producing a multi-subunit protein, comprising transforming a host cell with the expression vector of claim 1, and displaying the multi-subunit protein encoded by the vector onto the surface of the genetically replicable package.

30

40. The method of claim 39, wherein the vector comprises nucleotide sequences encoding functional portions of heterodimeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

41. A library of antibodies or antibody fragments made according to the method of claim 39.

42. A library of bacteriophage or phagemids, each carrying on its outer surface, one of a plurality of different-sequence polypeptides that comprises

- i. one of a plurality of first different-sequence heterologous polypeptide segments,
- ii. one of a plurality of a second different-sequence heterologous polypeptide segments,
- iii. joining the two segments, a peptide linker that has a cleavable peptide sequence that is not found in either of said polypeptide segments, and is recognized as a protein cleavage site by a proteolytic enzyme encountered in a bacteriophage host during bacteriophage biogenesis,

where cleavage of said linker by said host proteolytic enzyme results in a multimeric protein on the surface of a bacteriophage, each protein (i) having a plurality of different-sequence first and second polypeptides, and (ii) a protein activity related to the sequences of the first and second polypeptides.

43. The library of claim 42, wherein the protein activity is a specific binding affinity for a selected molecule of interest.

44. A library a bacteriophage genomes or phagemids, each genome encoding

- i. one of a plurality of first different-sequence heterologous polypeptide segments,
- ii. one of a plurality of a second different-sequence heterologous polypeptide segments,
- iii. joining the two segments, a peptide linker that has a cleavable peptide sequence that is not found in either of said polypeptide segments, and is recognized as a protein cleavage site by a proteolytic enzyme encountered in a bacteriophage host during bacteriophage biogenesis,

where cleavage of said linker by said host proteolytic enzyme results in a multimeric protein on the surface of a bacteriophage, each protein (i) having a plurality of

different-sequence first and second polypeptides, and (ii) a protein activity related to the sequences of the first and second polypeptides.

45. A method of identifying one or more multimeric proteins having a desired
5 above-threshold activity, comprising
producing a library a bacteriophage or phagemids, each carrying on its outer
surface, one of a plurality of different-sequence polypeptides that comprises
- i. one of a plurality of first different-sequence heterologous
polypeptide segments,
 - 10 ii. one of a plurality of a second different-sequence heterologous
polypeptide segments,
 - iii. joining the two segments, a peptide linker that has a cleavable
peptide sequence that is not found in either of said polypeptide
segments, and is recognized as a protein cleavage site by a
15 proteolytic enzyme encountered in a bacteriophage host during
bacteriophage biogenesis,
- where cleavage of said linker by said host proteolytic enzyme results in a
multimemric protein on the surface of a bacteriophage, each protein (i) having a plurality
of different-sequence first and second polypeptides, and (ii) a protein activity related to
20 the sequences of the first and second polypeptides, and
identifying bacteriophage in said library that have the above-threshold activity.

46. The method of claim 45, further comprising sequencing the portion of the
genome(s) of the identified bacteriophage that encode said first and second
25 polypeptides.

47. A method for creating a library of antibodies or antibody fragments,
comprising
- obtaining a biological sample;
 - 30 introducing the biological sample to a cell population capable of producing
antibodies;
 - reverse transcribing the light chain region and heavy chain region mRNA, or
fragments thereof, of the cell population;

amplifying and linking the two antibody fragment cDNA sequences with a linker comprising a nucleic acid sequence which encodes an amino acid sequence capable of being cleaved by a proteolytic agent;

5 amplifying the linked sequences to create a population of DNA fragments which encode the two antibody fragments;

cloning the population of DNA fragments into expression vectors and amplifying the cloned expression vectors;

10 selecting a subpopulation of expression vectors which encode antibodies or antibody fragments directed against the biological sample and amplifying the subpopulation selected to produce the library of antibodies or antibody fragments.

48. The method of claim 47, wherein said amplifying is performed by PCR.

15 49. A method for creating a patient-specific library of antibodies, comprising obtaining a sample of tissue from a patient; introducing the sample to a cell population capable of producing antibodies; reverse transcribing the light chain region and heavy chain region mRNA, or fragments thereof, of the cell population;

20 amplifying and linking the two antibody fragment cDNA sequences with a linker comprising an amino acid sequence capable of being cleaved by a proteolytic agent;

amplifying the linked sequences to create a population of DNA fragments which encode the two antibody fragments;

25 cloning the population of DNA fragments into expression vectors and selecting a subpopulation of expression vectors which encode recombinant anti-sample antibody fragments;

cloning the subpopulation of DNA fragments selected in-frame into expression vectors which encode antibody constant regions to produce intact antibody genes; and

expressing the subpopulation of intact antibody genes to produce the library of patient-specific antibodies.

30

50. An expression vector for expressing a multimeric polypeptide anchored on a surface of a genetically replicable package formed by a host, the expression vector comprising:

a vector segment encoding a polypeptide sequence having;

i. a first polypeptide segment having therein a first variable domain and a first constant domain of an antibody;

ii. a second polypeptide segment, and

5 iii. a third polypeptide segment having therein (a) a second variable domain and a second constant domain of an antibody, and (b) an anchoring peptide sequence for anchoring said multimeric polypeptide to said surface of said genetically replicable package,

10 the second polypeptide segment being between the first polypeptide segment and the third segment and having a length that prohibits the first and third polypeptide segments from associating intramolecularly to form a single-chain Fab, but allows two copies of the polypeptide to associate intermolecularly to form a di-Fab.

15 51. The expression vector of claim 50, wherein said second polypeptide segment further comprises a cleavable peptide sequence cleavable by a proteolytic agent.

20 52. The expression vector of claim 50, wherein the first polypeptide segment is N-terminal to the second polypeptide segment, and wherein the second polypeptide segment is N-terminal to the third polypeptide segment, and wherein the vector segment encoding the third polypeptide segment further includes one or more suppressable nonsense codon(s) N-terminal to the anchoring segment.

25 53. The expression vector of claim 50, wherein the third polypeptide segment further includes a cleavable peptide sequence cleavable by a proteolytic agent.

30 54. The expression vector of claim 51, wherein the proteolytic agent is selected from the group consisting of a chemical proteolytic agent and an enzymatic proteolytic agent.

55. The expression vector of claim 51, wherein the proteolytic agent is expressed by the host.

56. The expression vector of claim 51, wherein the proteolytic agent is added such that it contacts and cleaves the second polypeptide segment.

5 57. The expression vector of claim 54, wherein the chemical proteolytic agent is an acid.

58. The expression vector of claim 51, wherein the cleavable peptide sequence comprises the sequence represented by SEQ ID NO:1.
10

59. The expression vector of claim 51, wherein the cleavable peptide sequence is not found in either the first or third polypeptide segments, and is recognized as a protein cleavage site by a proteolytic agent encountered in the host.

15 60. The expression vector of claim 50, wherein the polypeptide sequence further comprises one or more leader sequence(s) positioned upstream of the first polypeptide segment or third polypeptide segment or both first and third polypeptide segments.

20 61. The expression vector of claim 50, wherein the anchoring peptide comprises a segment encoding a phage coat protein.

62. The expression vector of claim 50, wherein the expression vector is selected from the group consisting of plasmids, phages, cosmids, phagemids, and viral vectors.

25 63. The expression vector of claim 50, wherein the expression vector is selected from the group consisting of M13, f1, fd, If1, Ike, Xf, Pf1, Pf3, λ , T4, T7, P2, P4, ϕ X-174, MS2 and f2.

30 64. The expression vector of claim 50, wherein the genetically replicable package is selected from the group consisting of a bacteriophage, a virus, a cell and a spore.

65. The expression vector of claim 64, wherein the cell is a bacterial cell.

66. The expression vector of claim 65, wherein the bacterial cell is selected from the group consisting of strains of *Escherichia coli*, *Salmonella typhimurium*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Neisseria gonorrhoeae*, and *Bacillus subtilis*.

67. The expression vector of claim 64, wherein the cell is a yeast cell.

68. The expression vector of claim 50, wherein the genetically replicable package is a filamentous bacteriophage specific for *Escherichia coli* and the anchoring peptide is a phage coat protein selected from the group consisting of coat protein III, coat protein pVI and coat protein VIII.

69. The expression vector of claim 68, wherein the filamentous bacteriophage is selected from the group consisting of M13 and fd.

70. The expression vector of claim 51, wherein the proteolytic agent is encoded by a nucleic acid sequence in the expression vector.

71. The expression vector of claim 51, wherein the proteolytic agent is encoded by a nucleic acid sequence in a second expression vector.

72. The expression vector of claim 51, wherein the cleavable peptide sequence comprises a disordered region cleavable by the proteolytic agent.

73. The expression vector of claim 51, wherein the cleavable peptide sequence comprises a specific peptide cleavage site cleavable by the proteolytic agent.

74. The expression vector of claim 51, wherein the cleavable peptide sequence includes a cleavage site for urokinase, pro-urokinase, thrombin, enterokinase, plasmin, plasminogen, TGF- β , staphylokinase, thrombin, Factor IXa, Factor Xa, a metalloproteinase, an interstitial collagenase, a gelatinase or a stromelysin.

75. The expression vector of claim 51, wherein the cleavable peptide sequence is cleavable by a protease selected from the group consisting of degP, degQ, degS and tsp.

5 76. The expression vector of claim 51, wherein the cleavable peptide sequence comprises a self-cleaving domain.

77. The expression vector of claim 76, wherein the self-cleaving domain is derived from an intein.

10

78. A host cell comprising the expression vector of claim 50.

15

79. A method of producing a multi-subunit protein, comprising transforming a host cell with the expression vector of claim 50, and displaying the multi-subunit protein encoded by the vector onto the surface of the genetically replicable package.

20

80. A library of antibodies or antibody fragments made according to the method of claim 79.

81. A method of producing a di-Fab, the method comprising expressing the polypeptide sequence from the expression vector of claim 50 under conditions effective to allow the two copies of the polypeptide to associate intermolecularly to form a di-Fab.